Indian Journal of Chemistry
Vol. 62, December 2023, pp. 1252-1267
DOI: 10.56042/ijc.v62i12.5525

Virtual screening and molecular docking study of some naturally available phytochemicals against SARS-CoV-2

Avishek Dey[a], Iqrar Ahmad[b,c], Keshab Mondal[d], Rathin Jana[e], Harun Patel[b] & Soumen Mistri*[d]

[a] Department of Botany, Ramananda Centenary College, Laulara 723 151, Purulia, West Bengal, India

[b] Department of Pharmaceutical Chemistry, Prof. Ravindra Nikam College of Pharmacy, Gondur, Dhule 424 002, Maharashtra, India

[c] Division of Computer Aided Drug Design, Department of Pharmaceutical Chemistry, R. C. Patel Institute of Pharmaceutical Education and Research, Shirpur 425 405, Maharashtra, India

[d] Department of Chemistry, Ramananda Centenary College, Laulara 723 151, Purulia, West Bengal, India

[e] Department of Chemistry, Shahid Matangini Hazra Government General Degree College for Women, Nimtouri 721 649, Purba Medinipur, West Bengal, India

E-mail: soumen.mistri@gmail.com

Received 14 August 2023; accepted(revised) 24 November 2023

The novel human corona virus disease 2019, also known as COVID 19 or SARS-CoV-2 has designated as severe acute respiratory syndrome coronavirus 2, which has first emerged in Wuhan, China at the end of 2019. Now a day, it is a great challenge to scientists in the area of biology and chemistry to develop anticoronaviral drugs to overcome from this disease. SARS-CoV-2 was identified as a single-stranded positive-sense RNA virus. In this study, we have used virtual screening and molecular docking investigation of some naturally occurring bioactive organic compounds having the various phytochemical properties to compare the potential inhibitory activity of these molecules against SARS-CoV-2 protease. Based on ADME analysis and molecular docking study, Amentoflavone, Kazinol A, Kazinol B, Berbamine, Broussoflavan A, (-)-Catechingallate and Juglanin exhibits remarkable potentiality to bind with M^{pro} as compared to native ligand N3. Moreover the bioavailability radar study shows that Broussoflavan A is the only compound which is orally bioavailable among the studied compounds. A molecular dynamics simulation study of the ligand (Broussoflavan A) with protein indicates that the complex is stable. The docking study and MD simulation study indicates that in the protein-ligand complex, Broussoflavan A interact with active site of SARS-CoV-2 main protease.

Keywords: Naturally Available Phytochemicals, Virtual Screening, Molecular Docking, MD-Simulation

The entire world has been suffered due to emergence of the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) at the end of 2019 that cause life-threatening coronavirus disease 2019[1]. The disease was first reported in Wuhan, China and after that it spread rapidly through human to human transmission in the whole country[2] and consequently the World Health Organization (WHO) declared the disease as a pandemic on 11 March 2020[3-4]. Patients with COVID-19 have frequently suffered from acute respiratory distress that may lead to mortality due to substantial alveolar damage and eventually respiratory failure[5]. WHO reported on 2nd May 2020 that the virus had caused almost 3267184 infections and 229971 casualties all over the world. Therefore, a universal effort is badly needed to find out effective medicines especially antiviral drugs against this deadly disease. It becomes a great challenge to the scientists to develop anticoronaviral drugs to overcome from this disease, although some protease inhibitors and antiviral agents like oseltamivir[6], lopinavir[7], ritonavir[7], remdesivir[8], favipiravir[9], ribavirin[9], chloroquine[10] and hydroxychloroquine[10], have already been reported as potential therapeutic agents for COVID-19. The major goal of therapeutic research is to develop drug against SARS-CoV-2 to stop the life cycle of the virus by preventing host-virus interaction[11-12]. Moreover, several other protocols including chloroquine derivatives[13], azithromycin[14] and convalescent plasma[15] were also tested for therapeutic purpose.

SARS-CoV-2 was identified as a single-stranded positive-sense RNA virus and genetically belongs to betacoronavirus consisting of four major structural proteins, the spike protein, envelope protein, membrane protein, nucleocapsid protein and some

non-structural proteins (Nsp)[16-18]. Anand *et al*. reported that corona viruses produce short polypeptide chains during transcription by using proteolytic enzymes like papain-like protease (PLpro) and 3-chymotrypsin-like protease (3CLpro) to form a variety of non-structural proteins which are essential for viral replication[19]. In addition, unusual spike protein has also been discovered to have a high affinity for the human ACE2 receptor[20]. Therefore, it becomes crucial to focus on both primary protease and the spike protein to develop vaccines and other therapeutic agents against this virus.

Molecular docking is an important computational tool due to its ability to depict interactions between the ligand and its biological targets in the drug discovery domain[21,22]. These techniques support industry to focus on the compounds which have great potential against the specific target and also help in predicting the side and toxic effects of compounds[23]. Application of high-throughput protein purification techniques, crystallography, and nuclear magnetic resonance spectroscopy help us to study the structural details of fundamental corona virus enzyme SARS-CoV-2 main protease (Mpro) which plays a crucial role for mediating viral transcription and replication[24]. Mpro belongs to a protease class of proteins and this particular protease is seen to play essential role in replication process of many viruses[24]. Moreover, it was noted that the catalytic site is conversed in SARS-CoV-2 Mpro and it is made up of two amino acids including His41 and Cys145[25]. Literature survey shows that scientist established the crystal structure of N3 encapsulated COVID-19 virus (M^{pro}) (PDB ID: 6LU7) by using computer-aided drug design as well as virtual screening of more than 10,000 compounds[26].

Natural substances are well known for being abundant sources of antiviral drugs because of their wide range of structural variations and safety protocol. Since early days, plants have been used as a source of therapeutic phytochemicals and applied into various therapy regimens of viral diseases[27]. Natural products may have a significant role in supportive and prophylaxis management since treatment options for COVID-19, consists of only preventive and supportive procedures[28]. Literature survey shows that numerous dietary compounds, such as curcumin, savinin, and betulinic acid, have been found to inhibit SARS-CoV disease[29]. Moreover, a number of naturally available bioorganic compounds including kaempferol, quercetin, luteolin-7-glucoside, demethoxycurcumin, naringenin, apigenin-7-glucoside, oleuropein, curcumin, catechin, and epicatechin-gallate[30] have been shown to exhibit potential anti-COVID-19 agent using molecular docking investigation. In addition, it was noted that Forsythoside A exhibits antiviral effects against infectious bronchitisvirus (IBV), a virus that is a member of the coronavirus family[31].

Flavonoids, an important class of polyphenolic compounds found to be distributed widely throughout the plant kingdom[32] and has been classified into several subgroups like Chalcones, flavanes, flavanols, flavanones, flavanonols, flavones, flavonols, isoflavones or catechins, and procyanidins having a basic flavan (2 phenylchroman) structure in common. Flavonoids found to be associated with diverse therapeutic applications including prevention of the effects of reactive oxygen species[33], anti-cancer[34], antiviral, and anti-inflammatory[35]. Importantly, various flavonoids have been found to inhibit a variety of SARS-CoV targets by blocking enzymatic activities of viral proteases like 3-chymotrypsin like proteases (3CLpro), papain-like proteases (PLpro) and also interfering with the S protein[36]. Isobavachalcone (IBC) a naturally occurring chalcone blocks both corona viruses (SARS-CoV and MERS-CoV) that cause severe acute respiratory syndrome[37,38]. Moreover, it was noticed that flavonoids were connected with the suppression of the activity of angiotensin-converting enzyme (ACE)[39]. Furthermore, anti-inflammatory properties of flavonoids precisely activate and stimulate the host immune response during viral infections[40] as well as suppression of SARS-CoV-2 triggering overwhelming inflammatory reactions[41] were also investigated.

The present study is carried out to evaluate the effectiveness of some medicinal plant-based active chemicals like anthraquinone, biflavonoid, flavonoid, flavanone, alkaloid, triterpenoid, triterpene, chalcone, polyphenol, gallic acid ester, lectin, diphenylheptanoids, tannin, triterpenoidsaponin, flavonol, flavones and glycoside to compare the potential inhibitory activity against COVID-19 protease and also characterized the different properties of these compounds towards the SARS-CoV-2 main protease (Mpro) (in complex with the inhibitor N3) by molecular docking analysis. Application of phytochemicals as antiviral agents now

becomes a promising therapeutic approach globally because of less toxicity and minimum chance of developing resistance[42]. Therefore, the main focus of this work is to discover some potential plant-based antiviral agents against the Covid-19 virus using molecular docking study.

Experimental Section

Proteins/Macromolecules

COVID-19 $3cl^{pro}/M^{pro}$ (PDB ID: 6LU7) structures were obtained from https://www.rcsb.org/ in .pdb format. PDB is an archive for the crystal structures of biological macromolecules, worldwide[43]. The 6LU7 protein consists of two chains (A and B chain), which form a homodimer. We used both the chain for macromolecule preparation. The native ligand for 6LU7 is n-[(5-methylisoxazol-3-yl)carbonyl]alanyl-l-valyl-n~1~-((1r,2z)-4-(benzyloxy)-4-oxo-1-{[(3r)-2-oxopyrrolidin-3-yl]methyl}but-2-enyl)-l-leucinamide (N3).

Ligand and Drug Scan

The three dimensional (3D) structures of ligands were obtained from zinc (https://zinc.docking.org/), in .sdf format. ZINC is a public access database and tool set which is initially developed to enable ready access to compounds for virtual screening[44], that has become ever widely used for virtual screening[45-52], ligand discovery[53-56], pharamcophore screens[57], benchmarking[58-60], and force field development[61]. In this article, we have used several ligands which already have potential activity in several biomedical applications. The compounds used in the present study were Aloe emodin, Amentoflavone, Apigenin, Corylifol B, Kazinol A, Kazinol B, Kazinol F, Kazinol J, Luteolin, Mimulone, Myricetin, 3′-*O*-methyldiplacol, 4′-*O*-methyldiplacol, 3′-*O*-methyldiplacone, 4′-*O*-methyldiplacone, Bavachinin, Nymphaeol A, Berbamine, Cepharanthine, Lycorine, Betulonic acid, Betulinic acid, Broussochalcone A, Bavachalcone, Isobavachalcone, 4'-*O*-methylbavachalcone, Broussoflavan A, Chrysin, (-)-Catechingallate, Concanavalin, Curcumin, Arjunolic acid, Juglanin and Neobavaisoflavone.

For the initial screening of bioactive organic molecules as Drug-like properties, we used a web-based tool named Swiss ADME (https://www.swiss adme.ch/) using Lipinski's rule of five. According to the Lipinski's rule of five, the compounds qualify as a ligand must have the following characteristics: (i) molecular weight < 500 Da, (ii) lipophilicity i.e. Log P<5, (iii) consists less than 5 number of H-bond donors, and (iv) must have less than 10 number of H-bond acceptor groups[62,63]. It is to be noted that the compounds which shows more than two violations from Lipinski's rule of five, are ruled out from further docking study.

Molecular docking

We have used AutoDockVina for protein-ligand docking. We have optimized the downloaded structure of 6LU7 and each ligand and the optimized structure was used for the docking purpose. In order to prepare protein structure for docking, water molecules and the inhibitor N3 molecule was removed and added polar hydrogen bonds, Kollman charges and Gasteiger charges using Biodiscovery Studio and AutoDock Tools[64]. A grid box of x, y and z coordinates -10.729204, 12.417653 and 68.816122 respectively, was prepared around the binding site of the protein. The output file of protein-ligand docking consists of a number of conformations of each ligand with various binding affinity corresponding to various distance from rmsd as well as best mode rmsd. The conformation corresponding to zero various distance from rmsd and zero best mode rmsd was selected and the docked complex was converted to a 2D structure in order to examine the various interactions, including hydrogen bond, C-H···π, *etc*., of ligands with 6LU7 at binding site of 6LU7.

Bioavailability radar

Drug-likeliness of the naturally available bioactive organic molecules which have effective binding energy with the target M^{pro} as compared to the inhibitor N3 molecule were subjected to forming a bioavailability radar using the SwissADME tool (https ://www.swiss adme.ch/) with respect to six physiochemical properties. This bioavailability radar infers whether the drug molecule/molecules are orally bioavailable or not. The six physiochemical properties are solubility, size, polarity, lipophilicity, flexibility and saturation. The pink shaded region defines the optimal values of the 6 parameters and deviation from these parameters on a large scale is suggestive of the ligand not being orally bioavailable[65].

Molecular dynamics (MD) simulation

The all-atom MD simulation was run for 100 ns to establish the protein ligand complex (Broussoflavan

A-SARSCoV-2 Mpro) conformational stability and dynamic behaviour. The MD simulation was run on a Z4 HP workstation with the configuration Ubuntu 22.04.2 LTS 64-bit, Intel Xeon W-2245 @ 3.90 GHz, 8-Cores, CUDA 12, and NVIDIA RTX A4000 graphics processing unit, using the Schrodinger's Desmond 6.9.137 MD simulation programme (version 2022-1)[66-68]. The protein ligand complex was solvated in a truncated orthorhombic box using the SPC water model. Na+ and Cl- counter ions were introduced to make the system electrically neutral. By preserving the ionic strength at 0.15 M salt concentration, the physiological *p*H was maintained. The neutral system was well equilibrated and minimized by applying the steepest descent and the limited-memory Broyden-Fletcher-Goldfarb-Shanno (LBFGS) methods with a maximum cycle of 2000 and a convergence threshold of 1 kJmol^{-1} [69-71]. All atomistic MD simulations for 100 ns MD generation were performed at a constant pressure of 1 atm and a constant temperature of 300 K. The snapshots were taken every 100 ps and kept for each MD simulation production trajectory interaction study.

Results and discussion

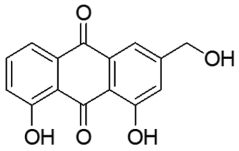
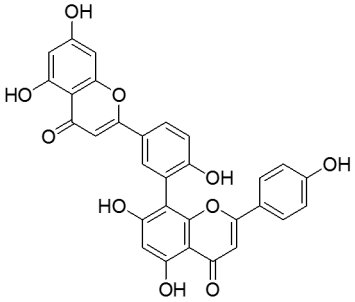
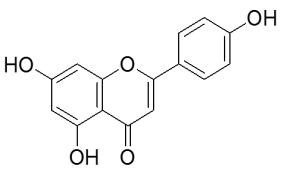
ADME analysis

Here we first randomly selected some naturally occurring bioactive organic compounds having the phytochemical properties like anthraquinone, biflavonoid, flavonoid, flavanone, alkaloid, triterpenoid, triterpene, chalcone, polyphenol, gallic acid ester, lectin, diphenylheptanoids, tannin, triterpenoidsaponin, flavonol, flavones and glycoside. Then we screen these compounds in order to find out their drug likeliness using Lipinski's rule of five. We select those bioactive organic compounds for molecular docking experiments with target protein which does not incur more than two violations of Lipinski's rule of five (Table 1).

Molecular docking

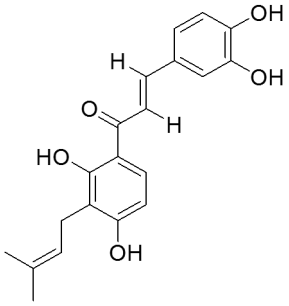
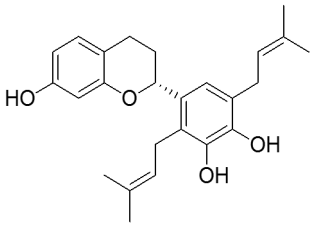
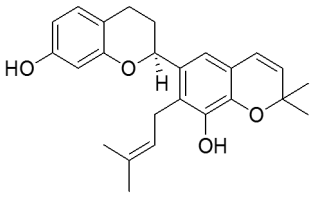
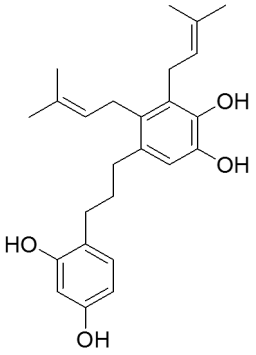
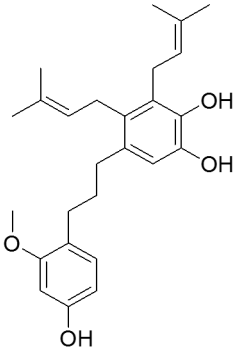
Molecular docking is an important computational tool in the drug discovery domain. It is done in order to find out the potential compounds which exhibit potential binding affinity to target protein and carry out the further study regarding bond formation in the protein-ligand complex at the binding site. In our

Table 1 — ADME analysis of 34 phytochemicals.

Sl. No.	Compound name	Compound structure	Zinc id	Lipinski's rule of five				
				Mw	logp	H-bond donor	H-bond acceptor	violations
Phytochemical Nature		Anthraquinone						
1	Aloe emodin		ZINC000004098644	270.24	1.50	3	5	0
Phytochemical Nature		Biflavonoid						
2	Amentoflavone		ZINC000003984030	538.46	3.62	6	10	2
Phytochemical Nature		Flavonoid						
3	Apigenin		ZINC000003871576	270.24	2.11	3	5	0

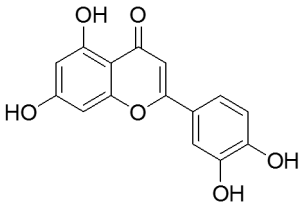
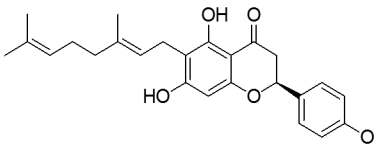
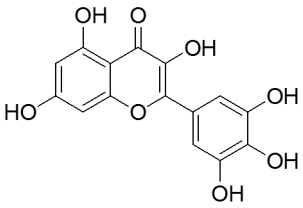
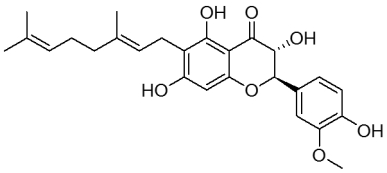
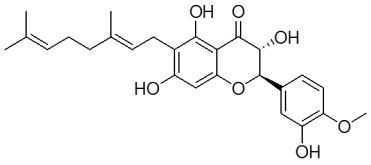
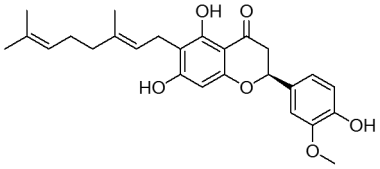
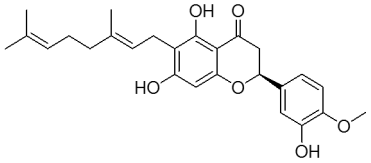
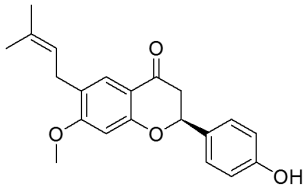
(Contd.)

Table 1 — ADME analysis of 34 phytochemicals. (*Contd*.)

Sl. No.	Compound name	Compound structure	Zinc id	Lipinski's rule of five				
				Mw	logp	H-bond donor	H-bond acceptor	violations
4	Corylifol B		ZINC000213285850	340.37	2.14	4	5	0
5	Kazinol A		ZINC000004098326	394.50	3.73	3	4	0
6	Kazinol B		ZINC000005854490	392.49	3.73	2	4	0
7	Kazinol F		ZINC000005158933	396.52	4.12	4	4	0
8	Kazinol J		ZINC000040872765	410.55	4.32	3	4	1

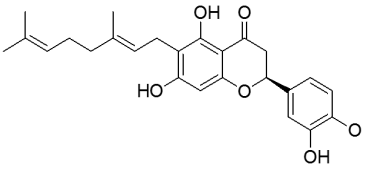
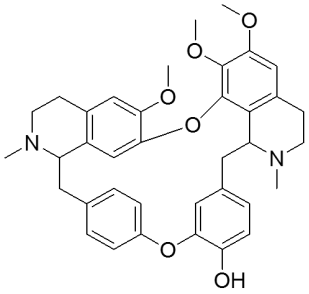
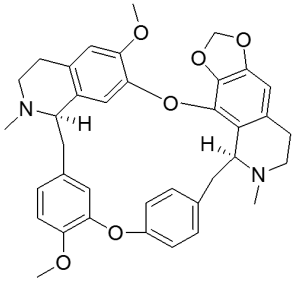
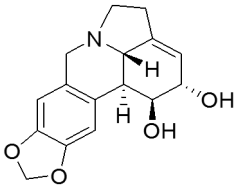
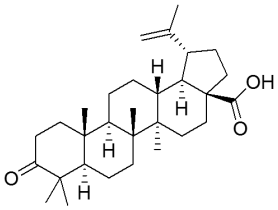
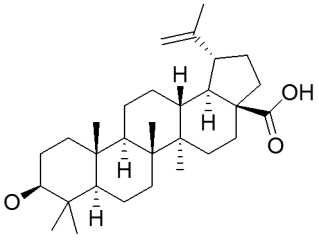
(Contd.)

Table 1 — ADME analysis of 34 phytochemicals. (*Contd*.)

Sl. No.	Compound name	Compound structure	Zinc id	Lipinski's rule of five				
				Mw	logp	H-bond donor	H-bond acceptor	violations
9	Luteolin		ZINC000018185774	286.24	–0.03	4	6	0
10	Mimulone		ZINC000013378636	408.49	2.82	3	5	0
11	Myricetin		ZINC000003874317	318.24	–1.08	6	8	1
12	3′-*O*-methyldiplacol		ZINC002356469714	454.51	1.68	4	7	0
13	4′-*O*-methyldiplacol		ZINC000015122018	454.51	1.68	4	7	0
14	3′-*O*-methyldiplacone		ZINC000015122023	438.51	2.48	3	6	0
15	4′-*O*-methyldiplacone		ZINC000015122025	438.51	2.48	3	6	0
Phytochemical Nature	Flavanone							
16	Bavachinin		ZINC000040862912	338.40	3.92	1	4	0

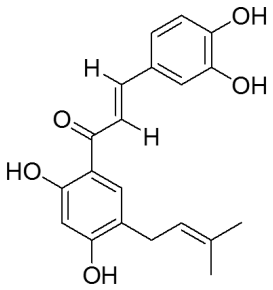
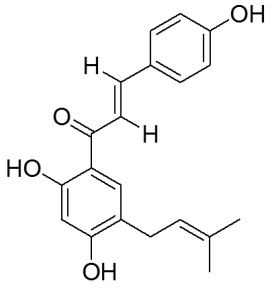
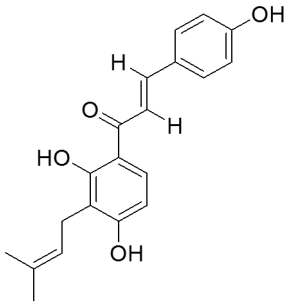
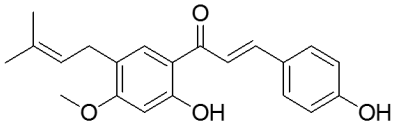
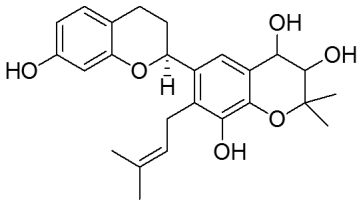
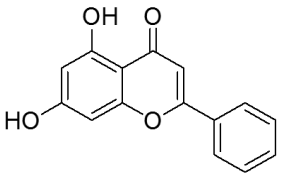
(Contd.)

Table 1 — ADME analysis of 34 phytochemicals. (*Contd*.)

Sl. No.	Compound name	Compound structure	Zinc id	Lipinski's rule of five				
				Mw	logp	H-bond donor	H-bond acceptor	violations
17	Nymphaeol A		ZINC000015122022	424.49	2.28	4	6	0
Phytochemical Nature		Alkaloid						
18	Berbamine		ZINC000030726840	608.72	5.13	1	8	1
19	Cepharanthine		ZINC000030726863	606.71	3.96	0	8	1
20	Lycorine		ZINC000003881372	287.31	1.08	2	5	0
Phytochemical Nature		Triterpenoid						
21	Betulonic acid		ZINC000008762252	454.68	5.73	1	3	1
Phytochemical Nature		Triterpene						
22	Betulinic acid		ZINC000118937400	456.70	5.82	2	3	1

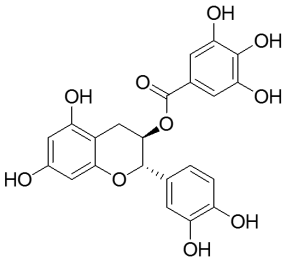
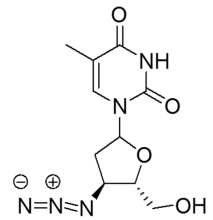
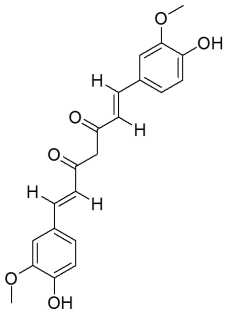
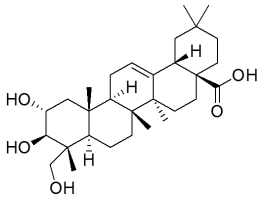
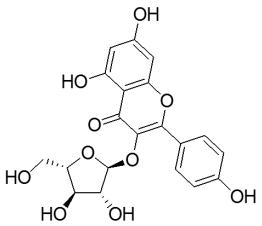
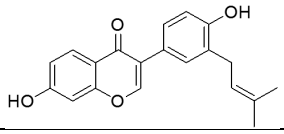
(Contd.)

Table 1 — ADME analysis of 34 phytochemicals. (*Contd*.)

Sl. No.	Compound name	Compound structure	Zinc id	Lipinski's rule of five				
				Mw	logp	H-bond donor	H-bond acceptor	violations
Phytochemical Nature	Chalcone							
23	Broussochalcone A		ZINC002356504112	340.37	2.14	4	5	0
24	Bavachalcone		ZINC002053598462	324.37	2.70	3	4	0
25	Isobavachalcone		ZINC000003925823	324.37	2.70	3	4	0
26	4'-*O*-methylbavachalcone		ZINC000019816069	338.40	2.92	2	4	0
Phytochemical Nature	Polyphenol							
27	Broussoflavan A		ZINC000040867311	426.50	2.16	4	6	0
28	Chrysin		ZINC000003872070	254.24	1.08	2	4	0

(Contd.)

Table 1 — ADME analysis of 34 phytochemicals. (*Contd*.)

Sl. No.	Compound name	Compound structure	Zinc id	Lipinski's rule of five				
				Mw	logp	H-bond donor	H-bond acceptor	violations
	Phytochemical Nature	Gallic acid ester						
29	(-)-Catechingallate		ZINC000004534390	442.37	0.32	7	10	1
	Phytochemical Nature	Lectin						
30	Concanavalin		ZINC000003779042	267.24	–1.25	2	7	0
	Phytochemical Nature	Diphenylheptanoid						
31	Curcumin		ZINC000000899824	368.38	1.47	2	6	0
	Phytochemical Nature	Triterpenoidsaponin						
32	Arjunolic acid		ZINC000031158038	488.70	4.14	4	5	0
	Phytochemical Nature	Flavonol						
33	Juglanin		ZINC000033832091	418.35	–1.57	6	10	1
	Phytochemical Nature	Flavone						
34	Neobavaisoflavone		ZINC000002570135	322.35	2.20	2	4	0

article, we choose 34 naturally available bioactive organic compounds from the ADME analysis and were subjected to molecular docking analysis against COVID-19 main protease of PDB id 6LU7. Fig 1 represents 3D structure of SARS-CoV-2 $M^{pro}$with native inhibitor N3, where ten amino acid residues (THR24, THR26, PHE140, ASN142, GLY143, CYS145, HIS163, HIS164, GLU166, HIS172) are present at the active site of the SARS-CoV-2 M^{pro}. The native inhibitor N3 was taken as a control and comparative study of the docking results of all 34 ligands (Table 2) with the control reveals that seven compounds [Amentoflavone, Kazinol A, Kazinol B, Berbamine, Broussoflavan A, (-)-Catechingallate and Juglanin] have remarkably better binding energy as compared with the binding energy of N3 (–7.20 kcal/mol) though other studied compounds have almost equivalent binding energy as N3. Table 2 shows the docking score as well as docking analysis results for several compounds against 6LU7, including ligand efficiency and inhibition constant.

Amentoflavone exhibits nine conformations at the binding site of 6LU7 and among the nine conformations, –9.90 kcal/mol was the least binding energy with corresponds to zero various distance from rmsd and zero best mode rmsd. In the best conformation, amentoflavone exhibits four different types of interactions including van der waals, H-bond, π···S and π···alkyl bond (Fig 2). SER46, TYR54, MET165 and ASP187 forming the conventional H-bond with amentoflavone, while MET49 and CYS145, MET 165 residues engaged with π···alkyl and π···S interactions, respectively, with amentoflavone. Other residue in Fig 2, involve in the weak van der Waals interaction with the ligand.

Kazinol A-M^{pro} complex shows minimum binding energy –8.40 kcal/mol and the interaction analysis of

Table 2 — Molecular docking analysis of several compounds against 6LU7.

	Binding Energy (kcal/mol)	Ligand Efficiency	Inhibition Constant
Aloe emodin	–7.20	–0.36	5.21×10^{-6}
Amentoflavone	–9.90	–0.25	5.43×10^{-8}
Apigenin	–7.70	–0.39	2.24×10^{-6}
Corylifol B	–7.10	–0.28	6.16×10^{-6}
Kazinol A	–8.40	–0.29	0.69×10^{-6}
Kazinol B	–8.40	–0.29	0.69×10^{-6}
Kazinol F	–7.30	–0.25	4.40×10^{-6}
Kazinol J	–7.30	–0.24	4.40×10^{-6}
Luteolin	–7.30	–0.35	4.40×10^{-6}
Mimulone	–6.40	–0.21	20.11×10^{-6}
Myricetin	–7.80	–0.34	1.88×10^{-6}
3′-*O*-methyldiplacol	–7.50	–0.23	3.13×10^{-6}
4′-*O*-methyldiplacol	–7.50	–0.23	3.13×10^{-6}
3′-*O*-methyldiplacone	–7.40	–0.23	3.71×10^{-6}
4′-*O*-methyldiplacone	–6.70	–0.21	12.11×10^{-6}
Bavachinin	–7.10	–0.28	6.16×10^{-6}
Nymphaeol A	–6.80	–0.22	10.23×10^{-6}
Berbamine	–9.60	–0.21	9.02×10^{-8}
Cepharanthine	–7.90	–0.18	1.59×10^{-6}
Lycorine	–7.50	–0.36	3.13×10^{-6}
Betulonic acid	–7.30	–0.22	4.40×10^{-6}
Betulinic acid	–7.50	–0.23	3.13×10^{-6}
Broussochalcone A	–6.90	–0.28	8.64×10^{-6}
Bavachalcone	–7.30	–0.30	4.40×10^{-6}
Isobavachalcone	–7.20	–0.30	5.21×10^{-6}
4'-*O*-methylbavachalcone	–7.00	–0.28	7.30×10^{-6}
Broussoflavan A	–8.70	–0.28	4.13×10^{-7}
Chrysin	–7.50	–0.39	3.13×10^{-6}
(–)-Catechingallate	–8.80	–0.28	3.49×10^{-7}
Concanavalin	–6.00	–0.32	39.54×10^{-6}
Curcumin	–7.00	–0.26	7.30×10^{-6}
Arjunolic acid	–8.00	–0.23	1.35×10^{-6}
Juglanin	–9.70	–0.32	7.62×10^{-8}
Neobavaisoflavone	–7.60	–0.40	2.65×10^{-6}

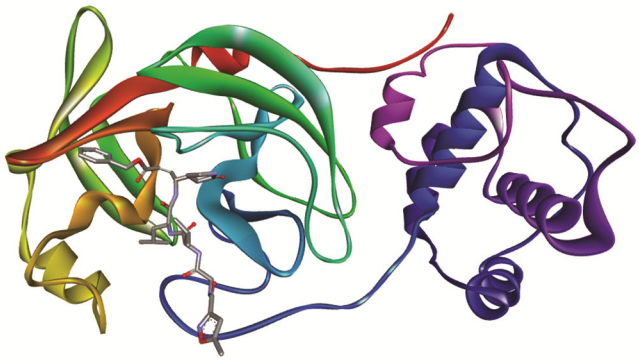


Fig. 1 — 3D structure of SARS-CoV-2 $M^{pro}$with native inhibitor N3.

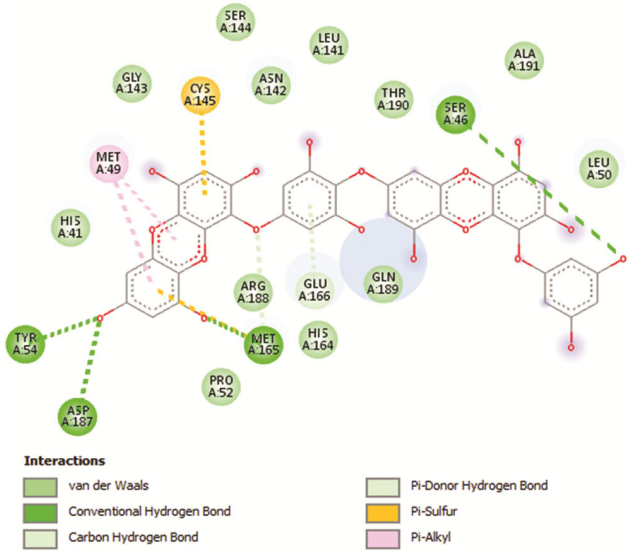


Fig. 2 — Interaction of Amentoflavone with 6LU7 protein.

the complex exhibits that Kazinol A forms the complex with M^{pro} *via* five different types of interactions (van der Waals, H-bond, alkyl, π···S and π···alkyl bond) (Fig 3). In the complex, LEU141, GLY143, SER144 and CYS145 residues engage in the formation of conventional hydrogen bond with Kazinol A whereas HIS41, MET49 residues forms alkyl bond with Kazinol A. It is also to be noted that MET49 and CYS145 residues also undergo π···alkyl and π···S interaction, respectively, with Kazinol A. Other residues in Kazinol A-M^{pro} complex engage in the van der Waals interaction with Kazinol A.

Another isomer of Kazinol, i.e., Kazinol B forms the complex with M^{pro} having the same minimum energy as Kazinol A and the interaction analysis of the complex exhibits three types of interactions including van der Waals, H-bond, and π···S bond. The minimum energy conformation of Kazinol B-M^{pro}complex exhibits that LEU141, ASN142, GLY143, SER144 and CYS 145 residues forms hydrogen bond with Kazinol B and the other residues as in Fig 4 involve in the van der Waals interaction. In addition, CYS 145 residue also undergo π···S interaction with Kazinol B in the minimum energy conformation of Kazinol B-M^{pro} complex.

Berbamine-M^{pro} complex shows minimum energy of –9.60 kcal/mol and in this conformation, the complex exhibits van der Waals, H-bond, π···anion, alkyl and π···alkyl interactions. In the minimum energy conformation of Berbamine-M^{pro} complex, MET165 and GLU166 residues involve in the hydrogen bond formation and π···anion interaction with Berbamine while LEU141 and MET165 residues involve in the formation of alkyl and π···alkyl bond, respectively (Fig 5).

Broussoflavan A forms the minimum energy complex with 6lu7 $M^{pro}$of energy –8.70 kcal/mol and in the complex SER144, CYS145 and MET165

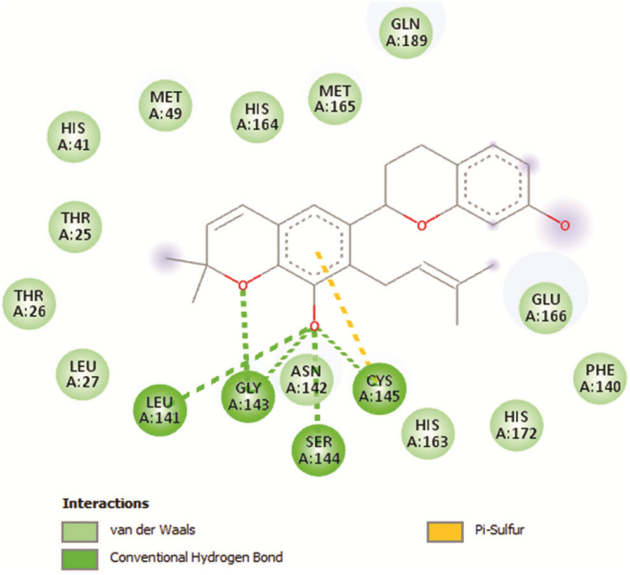


Fig. 4 — Interaction of Kazinol B with 6LU7 protein.

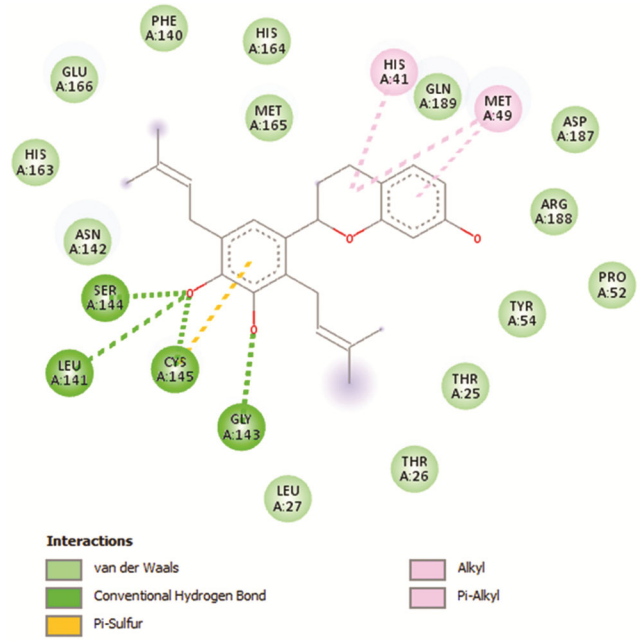


Fig. 3 — Interaction of Kazinol A with 6LU7 protein.

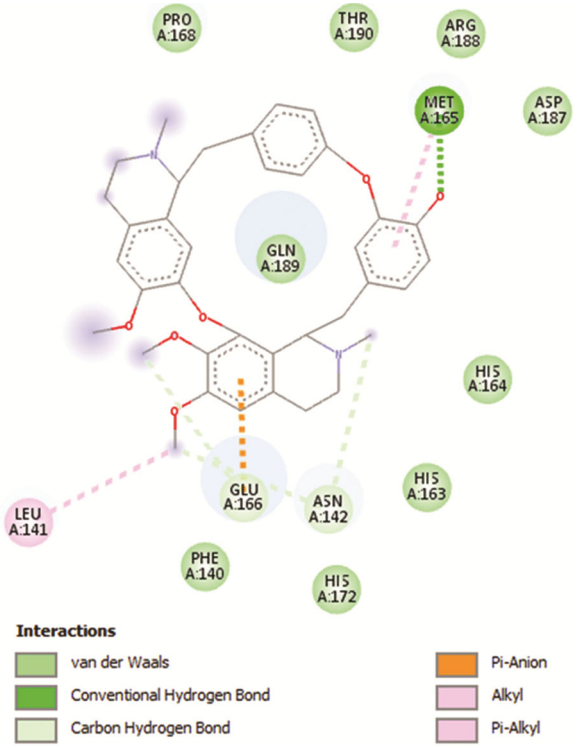


Fig. 5 — Interaction of Berbamine with 6LU7 protein.

residues forms hydrogen bond with Broussoflavan A while HIS41 and MET49 residues involve in alkyl and π···alkyl interaction (Fig 6). It is also to be noted that CYS145 residue also undergo π···S interaction with Broussoflavan A in the minimum energy conformation of Broussoflavan A- M^{pro}complex.

(–)-Catechingallate also forms complex with 6lu7 $M^{pro}$of minimum energy –8.80 kcal/mol and in the minimum energy conformation THR26, PHE140, GLU166, ASP 187 and MET49, CYS145 residues forms hydrogen and π···alkyl bond with (–)-Catechingallate, respectively. In addition, GLU166 also forms π···anion bond with (–)-Catechingallate in the minimum energy conformation of the complex (Fig 7).

Juglanin forms complex with 6lu7 Mpro of minimum energy –9.70 kcal/mol and the interaction analysis shows that LEU141, GLY143, SER144, CYS145, ASP187, THR190, GLN192 and MET49, CYS145 residue forms hydrogen bond and π···alkyl bond, respectively, with Juglanin where as HIS41 and MET165 undergo π··· π and π···S interaction, respectively (Fig 8).

M^{pro} belongs to a protease class of proteins which plays an important role in replication process of various viruses. 6LU7 was found as the main protease of SARS-CoV-2 in the pdb database and it is used first for the drug development process. There are huge structural similarities between the SARS-CoV-2 and SARS virus main protease. The catalytic site of SARS-CoV-2 main protease was conversed by two amino acids namely His41 and Cys145.

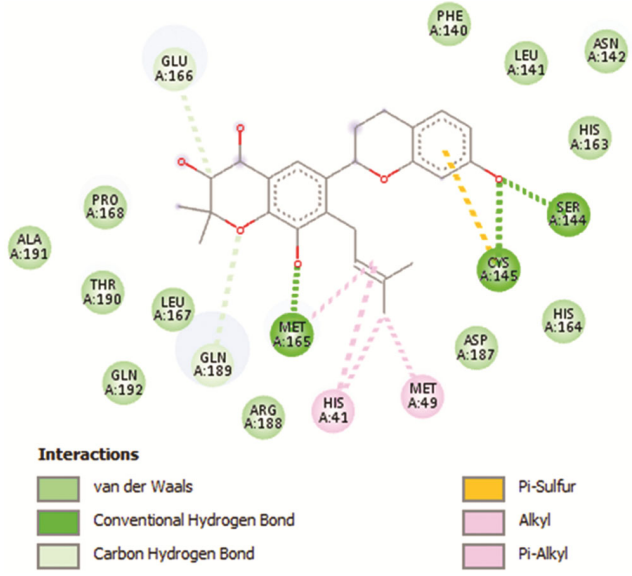


Fig. 6 — Interaction of Broussoflavan A with 6LU7 protein.

Bioavailability radar

The seven naturally available bioactive molecules namely Amentoflavone, Kazinol A, Kazinol B, Berbamine, Broussoflavan A, (–)-Catechingallate and Juglanin which have effectively more potentiality to bind with M^{pro} as compared to native ligand N3 were

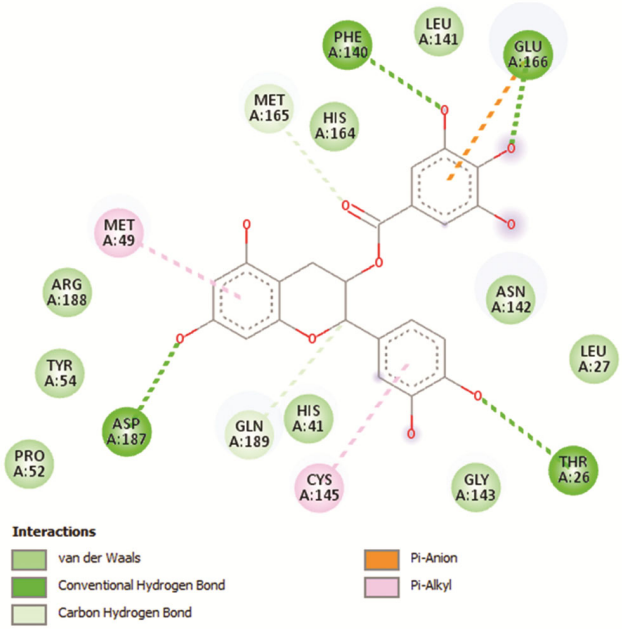


Fig. 7 — Interaction of (−)-Catechingallate with 6LU7 protein.

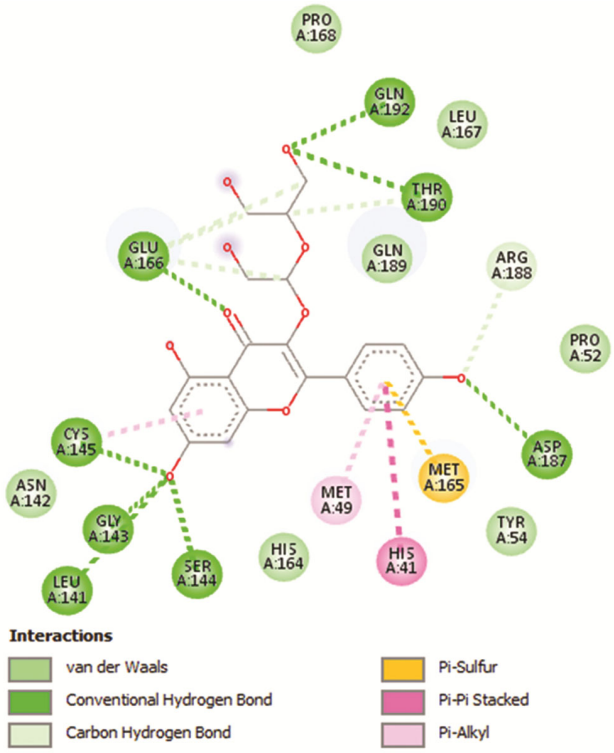


Fig. 8 — Interaction of Juglanin with 6LU7 protein.

subjected to bioavailability radar study which is a more comprehensive investigation to find out the drug-likeliness of ligand. Bioavailability radar is a descriptive tool to look into the drug-likeliness of the ligands based on six physiochemical properties. The present study indicates that Broussoflavan A to be orally bioavailable as the ligand radar entirely fit the pink shaded area (Fig 9). Kazinol A and Kazinol B have high value of lipophilicity (XLOGP3) as well as poor solubility [LogS(ESOL)] in water and consequently, these molecules are assigned as not orally bioavailable ligands (Fig 9). On the other hand, (–)-Catechingallate and Juglanin have relatively high polarity (TPSA) and so, these molecules also are no longer considered as orally bioavailable ligands (Fig 9). Amentoflavone is also not considered as orally bioavailable ligands, as it violate the lipophilicity (XLOGP3), size (MV), polarity (TPSA), solubility [LogS(ESOL)] and insaturation (fraction Csp3) parameter. In addition, Berbamine is not to be

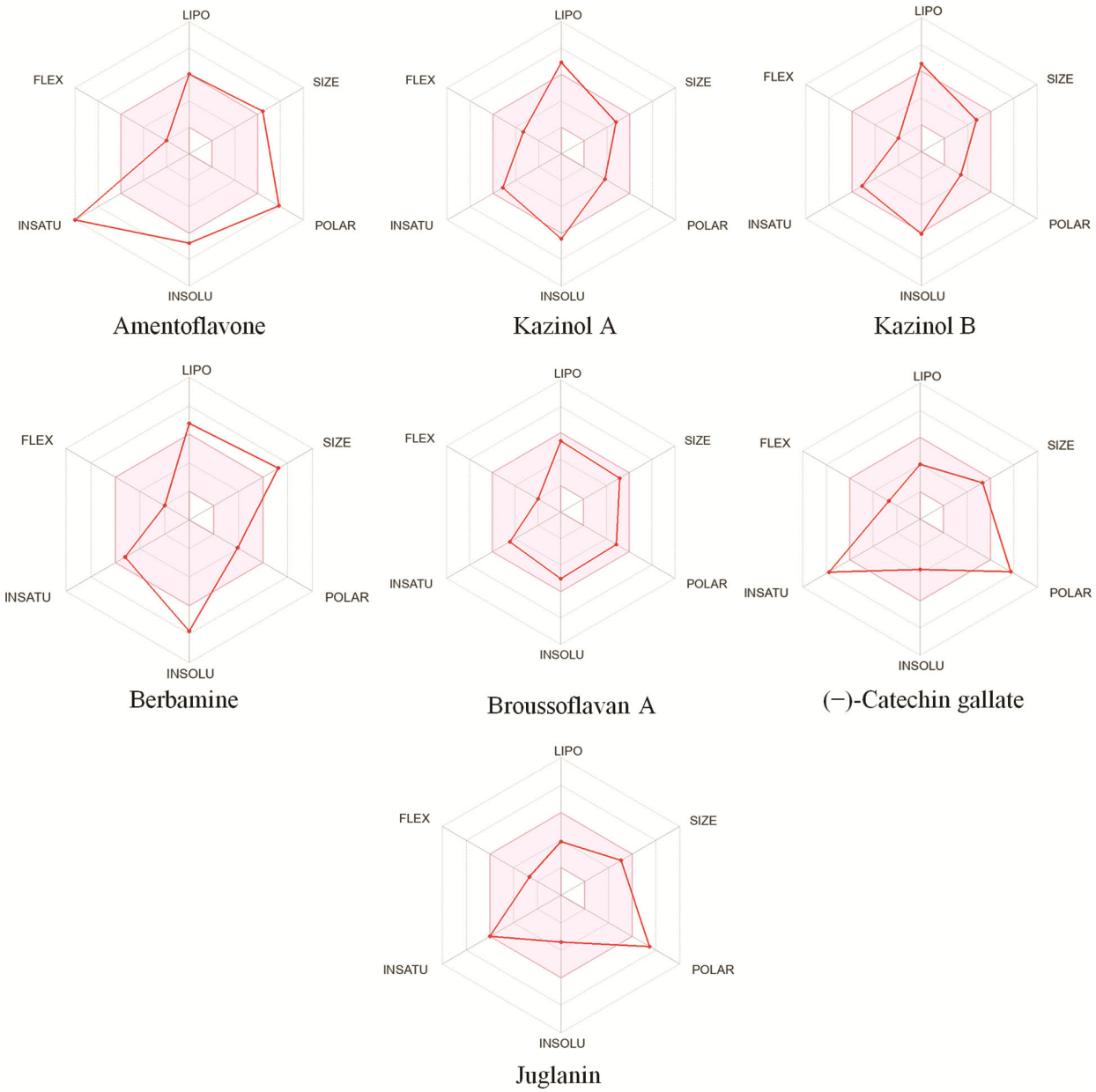


Fig. 9 — Bioavailability radar of Amentoflavone, Kazinol A, Kazinol B, Berbamine, Broussoflavan A, (−)-Catechingallate and Juglanin. The pink shaded zone is an estimated physicochemical space for oral bioavailability. LIPO (Lipophility): –0.7< XLOGP3<+ 5.0; SIZE: 150 g/mol<MV<500 g/mol. POLAR (Polarity): 20Å^2<TPSA<130Å^2. INSOLU (Insolubility): 0<LogS(ESOL)<6. INSATU (Insaturation): 0.25 < fraction Csp3<1. FLEX (Flexibility): 0<Number of rotatable bonds< 9.

considered as orally bioavailable ligands, as it disobey the size (MV) and solubility [LogS(ESOL)] parameter (Fig 9).

Molecular Dynamics (MD) simulation

A long range, up to 100 ns, MD simulation investigation was used to investigate the dynamic nature and binding/conformational stability of the Broussoflavan A- SARS CoV-2 Mpro complex. Several assessing parameters were investigated to study ligand-protein stability and dynamic behaviour during the simulation period, including the root mean square deviation (RMSD) of protein Cα atoms, ligand atoms, protein root mean square fluctuation (RMSF), and protein ligand contact mapping. The RMSD of each frame collected from the complete trajectory may be used to examine the structural conformation of the protein Cα atoms throughout the MD simulation. The little deviation and consistent change of the RMSD throughout the simulation demonstrate the protein-ligand complex's stability [72-73].

Fig 10A shows the RMSD of Cα atoms in each frame plotted against simulation length. Following ligand binding, the RMSD of protein Cα atoms was found to be efficiently equilibrated, with minute oscillations ranging from 0.9 to 2.40Å. The RMSD of the ligand was determined to be exceedingly low in comparison to the RMSD of the protein Cα atoms. The ligand RMSD was maintained at an average of 2.5Å after initial fluctuation due to equilibration. A converging pattern and minor RMSD fluctuations imply that the compound BroussoflavanA bound protein complex did not change structurally or conformationally during the simulation run.

The RMSF value is critical for determining the impact of individual amino acids on the stability of any protein-ligand interaction. It is the change in orientation of each amino acid Cα atom during the simulation as compared to the beginning orientation in the native state[74-76]. RMSF values were generated to investigate the effect of compound Broussoflavan A on the flexibility of SARS CoV-2 Mpro Cα atoms. From Fig 10B, it was observed that compound Broussoflavan A interacted with 21 amino acids of SARS CoV-2 Mpro protein during simulation which were His41(1.1 Å), Ser46(2.7 Å), Glu47(3.7 Å), Met49(2.3 Å), Lys97(2.9 Å), Asn142(2.1 Å), Ser144(1.5 Å), Cys145(1.1 Å), Gly146(0.8 Å), His163(0.8 Å), His164(0.7 Å), Met165(1.1 Å), Glu166(1.4 Å), Leu167(1.7 Å), Pro168(2.2 Å), Val186(1.4 Å), Arg188(1.7 Å), Gln189(2.0 Å), Thr190(2.2 Å), Ala191(2.3 Å), Gln192(2.0 Å). A vertical green bar identified all of these interacting residues. According to the RMSF graph, the variation in the SARS CoV-2 Mpro protein Cá atoms is an

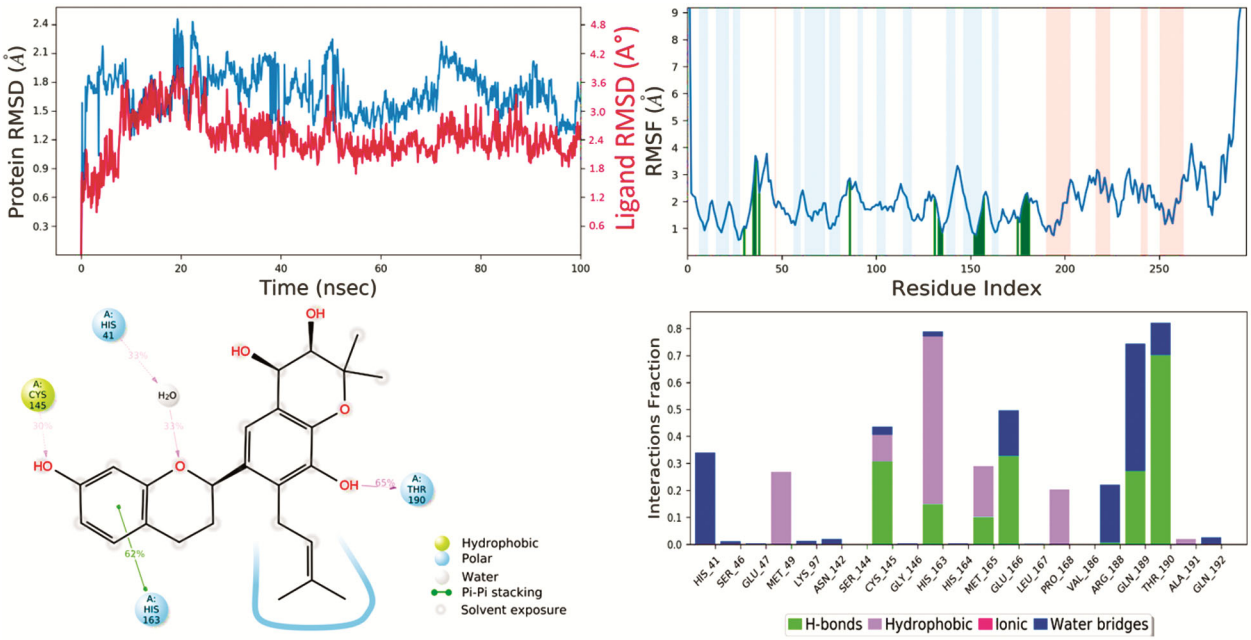


Fig 10 — MD simulation analysis of **Broussoflavan A** in complex with SARS-CoV-2 Mpro (PDB ID: 6LU7) (**A**) RMSD (Protein RMSD is shown in grey while RMSD of compound **Broussoflavan A** are shown in red) (**B**) Protein RMSF (**C**) 2d Interaction diagram and (**D**) Protein–ligand contact analysis of MD trajectory.

average RMSF value of 1.8 Å when it interacts with the compound Broussoflavan A.

The fluctuation can approach 3.0Å in certain loop locations (white colour backdrop), however it is much lower in the binding site region. The RMSF results showed that the SARS CoV-2 Mpro residues linked with compound Broussoflavan A remained constant throughout the simulation. The hydroxyl group of the chromane moiety established hydrogen bond contact with the critical residue catalytic dyad Cys145 at 30% of the simulation time in a 2D ligand interaction diagram from the 100 ns simulation, whereas residue Thr190 made hydrogen bond contact at 65%. At 62% of the simulation trajectory with His163, chromane developed hydrophobic π-π contacts with the phenyl group (Fig 10C). The molecule was located in the active pocket *via* creating hydrophobic contacts with the active site residues Met49, Cys145, His163, Met165, and Pro168, as well as water-mediated hydrogen bonding with His41, Glu166, Arg188, Gln189, and Thr190, according to the ligand interaction histogram (Fig 10D).

Conclusion

In this present work, virtual screening and molecular docking investigation of naturally available 34 phytochemicals against SARS-COV-2 have been carried out to detect the binding affinity of the phytochemicals at the catalytic site of the target protein. It was noticed that, seven naturally available bioactive molecules namely Amentoflavone, Kazinol A, Kazinol B, Berbamine, Broussoflavan A, (–)-Catechingallate and Juglanin exhibit remarkable potential in terms of binding at the catalytic site of the target protein as compared with native ligand N3. Among these seven bioactive molecules, only Broussoflavan A is orally bioavailable. Protein-ligand complex of Broussoflavan A exhibits interaction with His41 and Cis145 residues which are catalytic dyad in SARS-COV-2 M^{pro} in the docking study as well as in MD simulation study. Based on the present research observation and results, the phytochemical, Broussoflavan A should be precisely given high priority for development as prospective anti-COVID-19 remedy, though it requires vast pharmacological and clinical studies to understand its exact therapeutic values as anti-SARS-CoV-2 drugs.

Competing interests

The authors declare no competing interests.

Funding

Not applicable.

References

1 Wu Z &McGoogan J M, *JAMA*, 323 (2020) 1239.
2 Wang C, Horby P W, Hayden F G &Gao G F, *Lancet*, 395 (2020) 470.
3 Andersen K G, Rambaut A, Lipkin W I, Holmes E C & Garry R F, *Nature Medicine*, 26 (2020) 450.
4 Cucinotta D &Vanelli M, *Acta Bio-Med*,91 (2020) 157.
5 Xu Z, Shi L, Wang Y, Zhang J, Huang L, Zhang C, Liu S, Zhao P, Liu H, Zhu L, Tai Y, Bai C, Gao T, Song J, Xia P, Dong J, Zhao J & Wang F-S, *Lancet Respir Med*, 8 (2020) 420.
6 Li G &Clercq E D, *Nat Rev Drug Dis*, 19 (2020) 149.
7 Lim J, Jeon S, Shin H Y, Kim M J, Seong Y M, Lee W J, Choe K W, Kang Y M, Lee B & Park S J, *J Korean Med Sci,* 35 (2020) 1.
8 Holshue M L, DeBolt C, Lindquist S, Lofy K H, Wiesman J, Bruce H, Spitters C, Ericson K, Wilkerson S, Tural A, Diaz G, Cohn A, Fox L, Patel A, Gerber S I, Kim L, Tong S, Lu X, Lindstrom S, Pallansch M A, Weldon W C, Biggs H M, Uyeki T M &Pillai S K, *N Engl J Med*, 382 (2020) 929.
9 Wang M, Cao R, Zhang L, Yang X, Liu J, Xu M, Shi Z, Hu Z, Zhong W&Xiao G,*Cell Res* 30 (2020) 269.
10 Yao X, Ye F, Zhang M, Cui C, Huang B, Niu P, Liu X, Zhao L, Dong E, Song C, Zhan S, Lu R, Li H, Tan W&Liu D,*Clin Infect Dis*, 71 (2020) 732.
11 Cherrak S A, Merzouk H &Soulimane N M, *PLoS ONE*, 15 (2020) e0240653, (https://doi.org/10.1371/journal.pone.0240653).
12 Su H, Xu Y & Jiang H, *Fundamental Research,* 1 (2021) 151.
13 Gao J, Tian Z & Yang X, *BioSci Trends*, 14 (2020) 72.
14 Gautret P, Lagier J-C, Parola P, Hoang V T, Meddeb L, Mailhe M, Doudier B, Courjon J, Giordanengo V, Vieira V E, Dupont H T, Honoré S, Colson P, Chabrière E, Scola B L, Rolain J-M, Brouqui P&Raoult D, *Int J Antimicro Agents*, 56 (2020) 105949.
15 Chen L, Xiong J, Bao L & Shi Y, *Lancet Infect Dis*, 20 (2020) 398.
16 Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, Wang W, Song H, Huang B, Zhu N, Bi Y, Ma X, Zhan F, Wang L, Hu T, Zhou H, Hu Z, Zhou W, Zhao L, Chen J, Meng Y, Wang J, Lin Y, Yuan J, Xie Z, Ma J, Liu W J, Wang D, Xu W, Holmes E C, Gao G F, Wu G, Chen W, Shi W & Tan W, *Lancet*, 395 (2020) 565.
17 Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, Zhao X, Huang B, Shi W, Lu R, Niu P, Zhan F, Ma X, Wang D, Xu W, Wu G, Gao G F & Tan W, *N Engl J Med*, 382 (2020) 727.
18 Chen Y, Qianyun L &Guo D, *J Med Virol*, 92 (2020) 418.
19 Anand K, Ziebuhr J, Wadhwani P, Mesters J R&Hilgenfeld R, *Science*, 300 (2003) 1763.
20 Wan Y, Shang J, Graham R, Baric RS & Li F, *J Virol*, 94 (2020). (https://doi.org/10.1128/jvi.00127-20).
21 Kuntz I D, Blaney J M, Oatley S J, Langridge R &Ferrin T E, *J MolBiol*, 161 (1982) 269.
22 Kitchen D B, Decornez H, Furr J R &Bajorath J, *Nat Rev Drug Disc*, 3 (2004) 935.
23 Gao L, Xu J & Chen S, *Chin J Integr Med*, 26 (2020) 527.
24 Chen J, *Microbes Infect*, 22 (2020) 69.

25 Mirza M U &Froeyen M, *J Pharm Anal*, 10 (2020) 320.
26 Jin Z, Du X, Xu Y, Deng Y, Liu M, Zhao Y, Zhang B, Li X, Zhang L, Peng C, Duan Y, Yu J, Wang L, Yang K, Liu F, Jiang R, Yang X, You T, Liu X, Yang X, Bai F, Liu H, Liu X, Guddat LW, Xu W, Xiao G, Qin C, Shi Z, Jiang H, Rao Z & Yang H, *Nature*, 582 (2020) 289.
27 Lin S C, Ho C T, Chuo W H, Li S, Wang T T& Lin C C, *BMC Infect Dis*, 17 (2017) 144.
28 Solnier J &Fladerer J, *Phytochem Rev*, 20 (2021) 773.
29 Wen C, Kuo Y, Jan J, Liang P, Wang S, Liu H, Lee C, Chang S, Kuo C, Lee S, Hou C, Hsiao P, Chien S, Shyur L&Yang N, *J Med Chem*, 50 (2007) 4087.
30 Khaerunnisa S, Kurniawan H, Awaluddin R, Suhartati S &Soetjipto S, *Preprints,* (2020) (https://doi.org/10.20944/preprints202003.0226.v1).
31 Li H, Wu J, Zhang Z, Ma Y, Liao F, Zhang Y & Wu G, *Phytother Res*, 25 (2011) 338.
32 Wang T-Y, Li Q & Bi K-S, *Asian J Pharm Sci*, 13 (2018) 12.
33 Panche A N, Diwan A D & Chandra S R, J Nut Sci, 5 (2016) e47.
34 LeJeune T M, Tsui H Y, Parsons L B, Miller G E, Whitted C, Lynch K E, Ramsauer R E, Patel J U, Wyatt J E, Street D S, Adams C B, McPherson B, Tsui H M, Evans J A, Livesay C, Torrenegra R D & Palau V E, PLoS ONE, 10 (2015) e0142928.
35 Catarino M D, Talhi O, Rabahi A, Silva A M S & Cardoso S M, *Stud Nat Prod Chem*, 48 (2016) 65.
36 Yang Y, Islam M S, Wang J, Li Y & Chen X, *Int J BiolSci*, 16 (2020) 1708.
37 Jo S, Kim H, Kim S, Shin D H & Kim M S, *ChemBiol Drug Des*, 94 (2019) 2023.
38 Park J Y, Ko J A, Kim D W, Kim Y M, Kwon H J, Jeong H J, Kim C Y, Park K H, Lee W S &Ryu Y B, *J Enzyme Inhib Med Chem*, 31 (2016) 23.
39 Guerrero L, Castillo J, Quin˜ones M, Garcia-Vallve´ S, Arola L, Pujadas G &Muguerza B, PLoS ONE, 7 (2012) e49493.
40 Dong W, Wei X, Zhang F, Hao J, Huang F, Zhang C & Liang W, Sci Rep, 4 (2014) 7237. https://doi.org/10.1038/srep07237.
41 McKee DL, Sternberg A, Stange U, Laufer S &Naujokat C, Pharmacol Res, 157 (2020) 104859.
42 Ghildiyal R, Prakash V, Chaudhary V K, Gupta V & Gabrani R, *Plant-Derived Bioact*, (2020) 279, (https://doi.org/10.1007/978-981-15-1761-7_12).
43 Rosidi A, Khomsan A, Setiawan B, Riyadi H &Briawan D, *Pakistan J Nutr*, 15 (2016) 556.
44 Irwin J J&Shoichet B K, *J ChemInf Model*, 45 (2005) 177.
45 Irwin J J, Sterling T, Mysinger M M, Bolstad E S & Coleman R G, *J ChemInf Model*, 52 (2012) 1757.
46 Chen Y &Shoichet B K, *Nat ChemBiol*, 5 (2009) 358.
47 Carlsson J, Yoo L, Gao Z G, Irwin J J, Shoichet B K & Jacobson K A, J Med Chem, 53 (2010) 3748.
48 Naylor E, Arredouani A, Vasudevan S R, Lewis A M, Parkesh R, Mizote A, Rosen D, Thomas J M, Izumi M, Ganesan A, Galione A & Churchill G C, *Nat ChemBiol*, 5 (2009) 220.
49 Tikhonova I G, Sum C S, Neumann S, Engel S, Raaka B M, Costanzi S &Gershengorn M C, *J Med Chem*, 51 (2008) 625.
50 Irwin J J, Shoichet B K, Mysinger M M, Huang N, Colizzi F, Wassam P & Cao Y, *J Med Chem*, 52 (2009) 5712.
51 Leung W Y, Hamazaki T, Ostrov D A & Terada N, *J Mol Graphics Mod*, 45 (2013) 173.
52 Chen Y &Pohlhaus D T, *Drug Disc Today: Tech*, 7 (2010) e149.
53 Liu T, Lin Y, Wen X, Jorissen R N & Gilson M K, *Nucleic Acids Res*, 35 (2007) D198.
54 Bento A P, Gaulton A, Hersey A, Bellis L J, Chambers J, Davies M, Kruger F A, Light Y, Mak L, McGlinchey S, Nowotka M, Papadatos G, Santos R &Overington J P, Nucleic Acids Res, 42 (2014) D1083.
55 Zoete V, Cuendet M A, Grosdidier A &Michielin O, *J Comp Chem*, 32 (2011) 2359.
56 Blum L C &Reymond J L, *J Am ChemSoc*, 131 (2009) 8732.
57 Koes D R & Camacho C J, *Nucleic Acids Res*, 40 (2012) W409.
58 Mysinger M M & Shoichet B K, *J ChemInf Model*, 50 (2010) 1561.
59 Wishart D S, Knox C, Guo A C, Eisner R, Young N, | Gautam B, Hau D D, Psychogios N, Dong E, Bouatra S, Mandal R, Sinelnikov I, Xia J, Jia L, Cruz J A, Lim E, Sobsey C A, Shrivastava S, Huang P, Liu P, Fang L, Peng J, Fradette R, Cheng D, Tzur D, Clements M, Lewis A, DeSouza A, Zuniga A, Dawe M, Xiong Y, Clive D, Greiner R, Nazyrova A, Shaykhutdinov R, Li L, Vogel H J & Forsythe I, *Nucleic Acids Res*, 37 (2009) D603.
60 Kirchmair J, Markt P, Distinto S, Wolber G & Langer T, *J Comp.-Aided Mol Des*, 22 (2008) 213.
61 Sterling T & Irwin J J, *J ChemInf Model*, 55 (2015) 2324.
62 Lipinski C A, Lombardo F, Dominy B W & Feeney P J, *Adv Drug Deliv Rev*, 64 (2012) 4.
63 Giménez B G, Santos M S, Ferrarini M &Fernandes J P S, *Pharmazie*, 65 (2010) 148.
64 Morris G M, Huey R & Olson A J, *Curr Pro Bioinfo*, 24 (2008) 8.
65 Daina A, Michielin O &Zoete V, *Sci Rep*, 7 (2017) 42717.
66 Girase R, Ahmad I & Patel H, *J BiomolStructDyn*, (2023) 1, (https://doi.org/10.1080/07391102.2023.2203254).
67 Jagatap V R, Ahmad I, Sriram D, Kumari J, Adu D K, Ike B W, Ghai M, Ansari S A, Ansari I A, Wetchoua P O M, Karpoormath R & Patel H, *ACS Omega*, 8 (2023) 16228.
68 Sayed H M, Ramadan M A, Salem H H, Ahmad I, Patel H & Fayed M A, Steroids 196 (2023) 109245.
69 Zala A R, Rajani D P, Ahmad I, Patel H &Kumari P, *J BiomolStructDyn*, 41 (2023) 11518.
70 Osmaniye D, Ahmad I, Sağlık B N, Levent S, Patel H M, Ozkay Y &Kaplancıklı Z A, *J BiomolStrucDyn*, 41 (2022) 9022.
71 Kikiowo B, Ahmad I, Alade AA, Ijatuyi T T, Iwaloye O & Patel H M, *J BiomolStrucDyn*, 41 (2022) 10388.
72 Tabti K, Ahmad I, Zafar I, Sbai A, Maghat H, Bouachrine M &Lakhlifi T, *Comp BiolChem*, 104 (2023) 107855.
73 Benjamin I, Louis H, Ekpen F O, Gber T E, Gideon M E, Ahmad I, Ejiofor E U, *PolycycAromat Comp*, 43 (2022) 7942.
74 Khan M, Khan S, Alshammary F L, Zaidi S, Singh V, Ahmad I, Patel H, Gupta V K, Haque S, *J BiomolStrucDyn*, (2023), (https://doi.org/10.1080/07391102.2023.2209648).
75 Abdullahi M, Uzairu A, Shallangwa GA, Mamza PA, Ibrahim MT, Ahmad I & Patel H, J BiomolStructDyn, (2023), (https://doi.org/10.1080/07391102.2023.2208225).
76 Ayipo YO, Ahmad I, Chong CF, Zainurin NA, Najib SY, Patel H &Mordi MN, *J BiomolStrucDyn*, (2023), (https://doi.org/10.1080/07391102.2023.2198016).